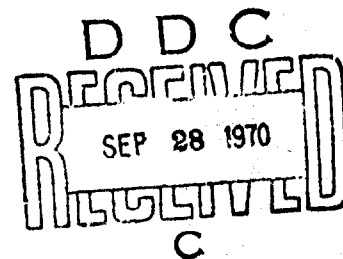


AD711759

EDUCATIONAL INFORMATION SYSTEM DESIGN
A CONCEPTUAL FRAMEWORK

F. W. Blackwell, B. W. Boehm,
A. Chalfant, J. A. Farquhar,
B. Markowitz, J. G. Root,
A. H. Rosenthal

August 1970



This document has been approved
for public release and sale; its
distribution is unlimited.

P-4411

EDUCATIONAL INFORMATION SYSTEM DESIGN[†]

A CONCEPTUAL FRAMEWORK

F. W. Blackwell, B. W. Boehm,
A. Chalfant, J. A. Farquhar,
B. Markowitz, J. G. Root,
A. H. Rosenthal

The RAND Corporation, Santa Monica, California

INTRODUCTION

The effectiveness of any educational institution greatly depends upon information collection and transfer. A variety of informational "inputs and outputs" are associated with each facet of the educational process. Many may be classified as routine, such as attendance and grade reporting; others, such as parent and teacher expectations, are highly subjective, yet equally important in judging the benefits of education. The term *information system* describes the full range of processes, methods, and techniques through which educational data are collected, permuted, and dispersed.

A multiplicity of alternatives confronts the administrator who considers building a formal information system. The complexity of such a system can range from simple manual records in classrooms and administrative offices to sophisticated automated systems serving large metropolitan areas. The administrator's needs may best be met by using a remote, already existing service, such as the California Educational Information System. Few "rules" exist for the administrator who has decided upon an information system as a solution to his problems; few relationships or measures of effectiveness for various

[†]Any views expressed in this Paper are those of the authors. They should not be interpreted as reflecting the views of The RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The RAND Corporation as a courtesy to members of its staff.

configurations have been established. Information system design is primarily an *ad hoc* activity, depending more upon expertise than on formal methodology.

A number of critical aspects of educational information system design are discussed below. The success of any information system depends largely upon several key decisions that must be made early in the preliminary design phase. This Paper attempts to (1) indicate the nature and some of the major implications of these key decisions, and (2) provide a framework for rational decisionmaking.

THE INFORMATION SYSTEM AND THE OPERATIONAL SYSTEM

An information system cannot be designed in the abstract, that is, without reference to the particular educational environment under scrutiny. (For convenience, we call this environment--the complex of curricula, standards, programs, and material and human resources--the *operational system*.) Conversely, some decisions on information-system implementation will have strong and direct effects on the operational system. A properly thorough treatment of information-system alternatives must cover all such interactions with the operational system. Table 1 shows the relationship among key information-system and operational-system decisions. The Paper is organized along the same lines as Table 1, that is, the section headings roughly parallel the boxes in the table. This organization conveniently illustrates the interdependence of pertinent design decisions. All too often, decisions are made by those responsible for one type of system without proper consideration of the full implications of the decision upon other areas. Thus, this table affords the administrator an indication of points where greatest care must be taken. In addition, it may serve as a mechanism to judge when "danger signals" are appearing in the domain of others.

Table 1

RELATIONSHIPS OF KEY INFORMATION SYSTEM AND OPERATIONAL SYSTEM DECISIONS

Decision	Affects Primarily Operational System	Affects Both Information and Operational Systems	Affects Primarily Information System
Operational-System Decisions	***1. Issues concerned with curricula, innovative programs, bureaucratic structure, etc.	***2. Overall time schedule of information system development. ***3. Relative level of expenditure on the information system. ***4. Educational Program Structure. ***5. Administrative structure. ***6. Level of accountability.	**7. Extent to which precautions should be taken to safeguard information privacy. *8. Extent of collection of follow-up data on graduates. *9. Provisions needed for correlations with external sociological data.
Information-System Decisions	**10. Amount of data-collection load to be placed on teachers and students. *11. Extent to which information systems will be used for research.	*12. Questions concerning information availability; quantity, level of detail, speed, and accuracy. *13. Extent to which information system should be concentrated at a single center.	***14. Choice of hardware configuration, information request monitoring, and data structure. *15. Whether or not to build a prototype information system.

Two factors somewhat offset the benefits of this approach. First, such questions seldom fit neatly into any single box; and second, the *relative* importance of the questions is sometimes obscured. In the first case, such assignments are not excessively misleading when viewed with a healthy skepticism. To minimize the second drawback, we indicate the relative importance of the points in Table 1 as follows:

- *** Decisions vital to the system's effectiveness.
- ** Decisions that can strongly affect the system's effectiveness.
- * Decisions that can affect several aspects of the system's effectiveness.

This classification of issues is used throughout the Paper.

OPERATIONAL-SYSTEM DECISIONS PRIMARILY AFFECTING THE OPERATIONAL SYSTEM

Issues Concerned with Curricula, Innovative Programs, Bureaucratic Structure, etc. [***]

The form of the information system must reflect the structure and purpose of the educational institution. This category is included as a summation of this point. Further discussion will be found under each of the headings enumerated below, where the various decisions are individually identified and examined.

OPERATIONAL-SYSTEM DECISIONS AFFECTING BOTH INFORMATION- AND OPERATIONAL-SYSTEMS

Overall Time Schedule of Information System Development [***]

Given an important project and strong financial backing for its early phases, a strong temptation often arises to achieve a working operational system as soon as possible.

This frequently leads to a situation in which the schedule for designing the associated information system heavily overlaps the schedule for implementing it; for example, beginning the layout of the data-collection forms before completing the study of desired processing outputs.

Unless the information system is strikingly similar to one that the designers have already implemented, simultaneous design and implementation will almost certainly produce revised specifications that lead to costly, time-consuming changes in the partially implemented forms of software. Well-documented examples of this phenomenon in educational information systems are difficult to find, but consider the published experience in another software field: producing FORTRAN compilers for different computers.[†]

Effort No. (same group)	Estimated Man-months	Actual Man-months
1	36	72
2	24	36
3	12	14

The successive ratios of actual development time and of estimation accuracy are fairly typical of information systems in general, whether or not they are computer-based.[‡]

[†]E. Naur and B. Randell (eds.), *Software Engineering*, NATO Scientific Affairs Division, Brussels, Belgium, January 1969, p. 74.

[‡]At first glance, there is little apparent similarity between sophisticated FORTRAN compilers and applications programs. However, despite the difference in program sophistication, the requirements of a FORTRAN compiler are rigidly defined by two constants (the structure of the FORTRAN language and the set of machine instructions). An educational information system has no such boundaries. Thus, it is much more difficult to define its realm, and to avoid redefinition during implementation. These considerations make estimating extremely difficult.

The most severe time limitations arise when one wishes an information system to monitor an experimental program. In many cases, there is no time to undertake an orderly development prior to the start of the program. However, there are other alternatives besides using an inappropriate existing information system or delaying experimentation until the new information system arrives. Perhaps the most attractive is the temporary use of an existing information system while a new system is being designed and implemented. System design could then be based upon the lessons learned through the (relatively) inexpensive experience with the pre-existing system. The price is a possible loss in completeness of longitudinal data. However, this procedure allows both an early start and orderly development of the information system, with the added advantage of being based on actual experience.

Relative Level of Expenditure on the Information System [***]

The appropriate level of expenditure for the information system is a difficult question, answerable only when the resources, problems, and requirements of the specific school are well understood. Obviously, the other points discussed in Table 1 will directly bear on the cost of the information system. Information system configuration (and attendant costs) must reflect many aspects of school operation. In particular, the amount of experimental and innovative activity materially ~~determines~~ the requisite scope and sophistication of the information system. A properly designed information system can greatly assist in certain areas of evaluation, as well as be responsive to the day-to-day needs of the school.

Several particulars on the cost and effectiveness of information systems deserve special consideration. As seen in Fig. 1, one buys information-system effectiveness with information-processing power, i.e., the capability to do a

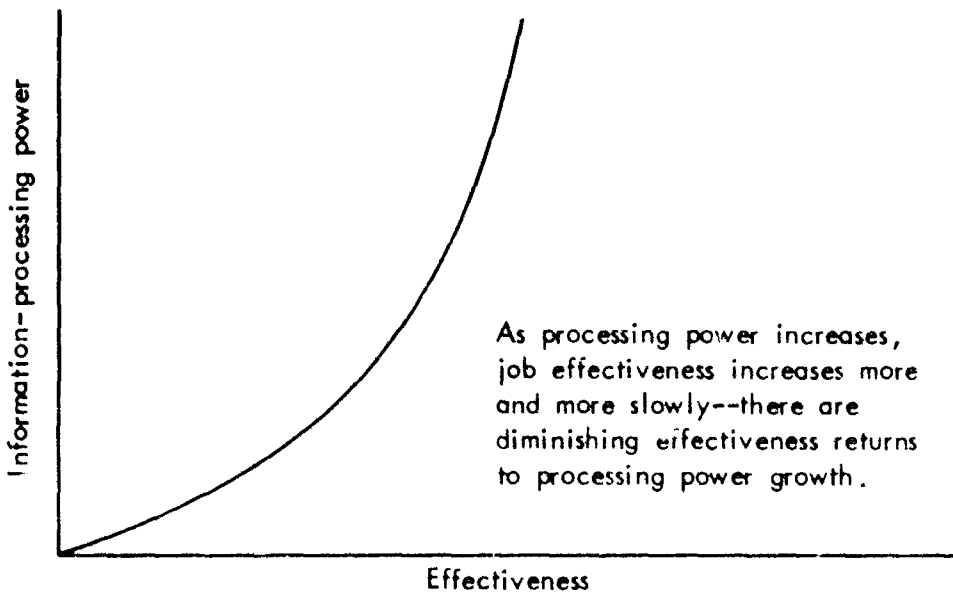
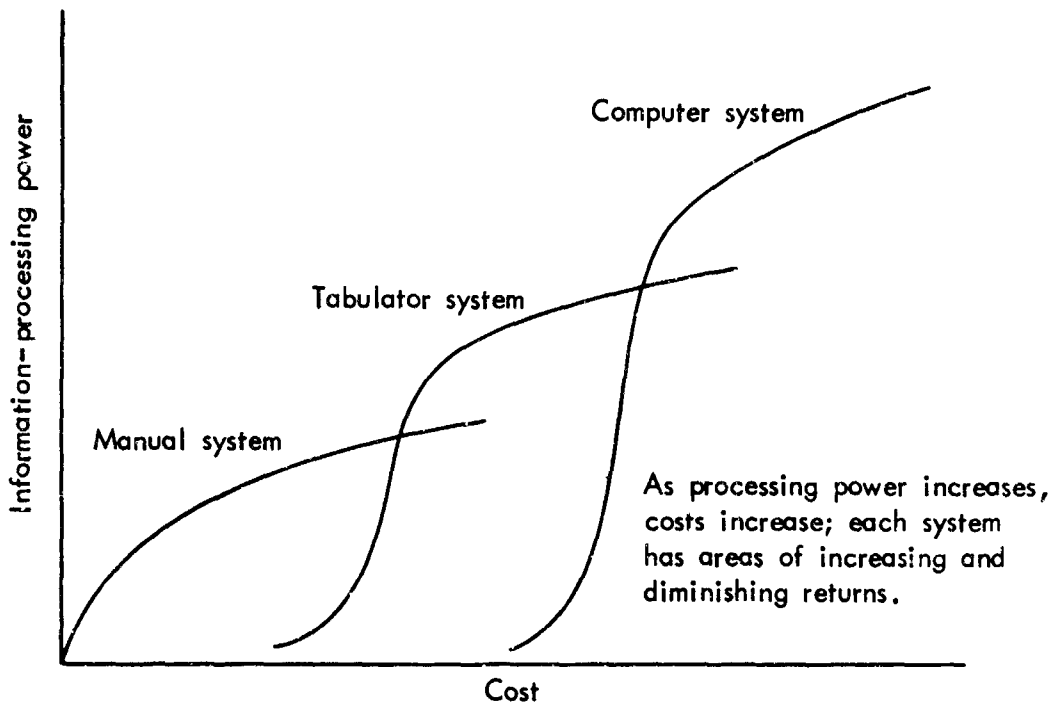


Fig. 1--Relationships Between Cost and Effectiveness of Information Systems

certain number of arithmetic and logical operations per hour. This generally relates to effectiveness in the smooth manner indicated in the lower (effectiveness) curve. However, the numerical values along this curve, especially that of the "point of diminishing returns," are highly dependent upon the character of programs and routine operations, and can be determined only after a good deal of analysis.[†] As the "plateau" characteristic of the set of upper (cost) curves indicates, a premature decision on a particular type of information-processing system (manual card-file, punch-card, and tabulation machines; or a computer-based system) may commit one to a higher or lower plateau of information-processing power than is necessary. At that point, re-designing and converting to the appropriate system can be expensive in both time and dollars.

Figure 2 shows a fairly typical breakdown and time-phasing of the expenditure rate for a relatively large educational information system.[‡] Note that only one-third of the annual operating cost is allocated for equipment (of the remaining two-thirds, 44 percent goes for staff salaries and 22 percent for operating costs), and that staffing levels for planning and design are comparable to those required during the operational phase. These figures are comparable to those reported for information systems in large corporations.^{††}

[†]It is important to understand that Fig. 1 is an expression of broad, imperfectly understood relationships. Any effort to assign numbers to either axis would require a detailed treatment far beyond the scope of this Paper. This is particularly true of the effectiveness scale. To our knowledge, no single effectiveness criterion exists that allows rational, quantitative comparison of manual, tab, and computer systems.

[‡]L. E. Yarbrough, "The Florida Plan for a Statewide Education Management Information System," *J. Assoc. Educational Data Systems*, September 1969, pp. 1-9.

^{††}W. F. Sharp, *The Economics of Computers*, The RAND Corporation, R-463-PR, August 1969. (Also, Columbia University Press, New York, 1969.)

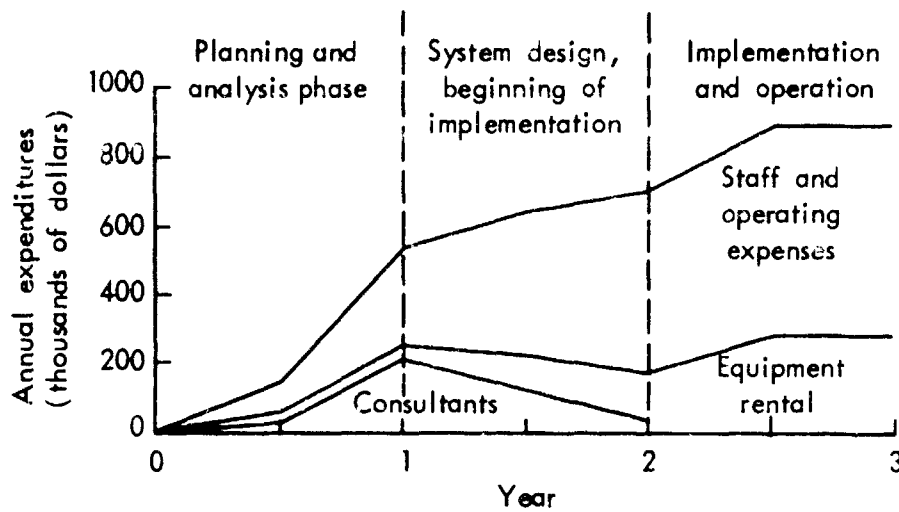


Fig. 2--A Sample Phased Budget for an Educational Information System

Some monthly rental ranges for two typical computer-based information systems are:[†]

Characteristics	Basic Card System	Basic Card-Tape System
Computer memory, words	4000	8000
Card reader, words/min	800-1000	800-1000
Card punch, words/min	250-300	250-300
Line printer, lines/min	700-1000	700-1000
Tape units	0	4
Monthly rental, dollars	700-1500	1500-5000

Of course, alternative ways of securing computing power (e.g., service bureaus and time-sharing systems) must also be evaluated, and other options considered (e.g., using a computer for both data management and computer-aided instruction).

[†]Keydata Corporation, *Computer Characteristics Review*, Vol. 9, No. 2, 1969.

Educational Program Structure [***]

An effective information system design must reflect the school's educational structure and philosophy. If many of the educational programs are upgraded, incorporating a large volume of grade-analytical routines would be wasteful. On the other hand, such an environment might well demand that the information system effectively store and retrieve qualitative and non-numerical information.

Administrative Structure [***]

Similar consideration must be made of the various possible aggregations. A school organized along "departmental" lines might effectively utilize reports of overall department activities and achievement, whereas other schools might prefer primary aggregation of data by grade, or based upon some criteria of measured ability. The information system should be designed so that a minimum of manual (or specially-requested computer) recalculation is necessary to fit the data to the existing bureaucratic and pedagogic structures.

A final relevant structural point concerns the context of the institution using the information system. In most cases, this institution will be required to submit reports to district, state, and even national authorities. Proper and insightful system planning can insure that this is done with a minimum of difficulty; poor planning can insure that it will be a painful, disruptive, and expensive experience.

Level of Accountability [***]

"Accountability" is a vogueish, if somewhat obscure (in operational implications) concept. It implies measurement of transactions involved in education: transactions both of human and material resources, and of learning itself. An information system can materially aid such a measurement

of the effectiveness of this transaction, and the identification of responsibility for it. The *level of accountability*--classroom, department, administrative--in turn defines several aspects of the information system (e.g., data collection mechanisms, and report structures). In addition, accountability demands integration of educational and resource data--a difficult part of system design.

OPERATIONAL-SYSTEM DECISIONS PRIMARILY AFFECTING INFORMATION-SYSTEM PERFORMANCE

Extent to Which Precautions Must be Taken to Safeguard Information Privacy [**]

A large file of organized information on student and teacher performance, perhaps including socioeconomic data and indicators of parent and teacher attitudes, presents a strong temptation for possible abuse for private profit or unauthorized use by public agencies.

Such safeguards as user passwords and file-protection programs, if designed into the system from the beginning, provide adequate security at an incremental information-system cost of just a few percent (to accommodate the extra routines, data fields, and validation required). However, adding the safeguards after system implementation adds from 25 to 50 percent to the information-system cost. Information privacy considerations have been discussed in detail elsewhere.[†]

Extent of Collection of Follow-Up Data on Graduates [*]

This concerns the collection of follow-up data on graduates of both the educational system and particular programs. The root of the problem is that geographical distribution expands as time increases. There is no easy solution;

[†]W. H. Ware, *Security and Privacy in Computer Systems*, The RAND Corporation, P-3544, April 1967.

the problem is mentioned here simply to indicate that the school in question should consider some commonality of system design, or at least lines of effective communication. This consideration is particularly appropriate within the State of California, where the California Educational Information System (CEIS) provides some measure of consistency.

Provisions Needed for Correlations with External Sociological Data [*]

If the implications of performing correlations--resolving differences between boundaries of school districts and census tracts, determining comparable data accuracy, providing cross-references between files--are thought through in advance, and the information system structured with respect to these considerations, then the job will be relatively straightforward. However, it may be quite difficult to provide for such correlations after the information system has been implemented.

INFORMATION-SYSTEM DECISIONS PRIMARILY AFFECTING OPERATIONAL-SYSTEM PERFORMANCE

Amount of Data-Collection Load to be Placed on Teachers and Students [**]

In implementing an information system, decisions could be made on data collection that might seriously impair teacher or student performance. For example, in an experimental situation, having the teacher monitor and record the class attention level during a programmed instruction sequence both reduces teaching effectiveness and makes the class atypical with respect to normal teacher performance.

Extent to Which Information System Will Be Used for Research [**]

A well-designed information system should be quite useful to a broad spectrum of potential "clients": teachers,

counselors, school and district administrators, parents. Furthermore, extending the information system could significantly increase its utility in the performance of research. The concept of such a "multipurpose-multiuser" information system has been explored in some detail by Coleman and Karweit and holds considerable promise for improving the effectiveness of the school system.[†] The cost of such efforts should be carefully evaluated. Design of an information system that would be "all things to all men" is probably not possible. In any event, the attempt would require massive expenditure. In that all research requires different data, the tradeoffs between flexibility, comprehensiveness, and cost must be carefully examined.

INFORMATION-SYSTEM DECISIONS AFFECTING BOTH INFORMATION- AND OPERATIONAL-SYSTEMS

Questions Concerning Information Availability: Quantity, Level of Detail, Speed, and Accuracy [***]

Questions concerning the speed, accuracy, and quantity of information that should be made available for requests are fundamental to the design of the information system; their answers determine the goals of the hardware-software design process (see p. 15 below). Formulating the answers requires close collaboration between information recipients and information-system designers. However, a further key decision must be made on the extent to which the information system will support educational research along the "multi-purpose-multiuser" lines envisioned by Coleman and Karweit.

These considerations provide the guidelines for the extent, level of detail, and retrievability of the information

[†] J. S. Coleman and N. L. Karweit, *Multi-Level Information Systems in Education*, The RAND Corporation, P-4377, June 1970.

stored in the data base. For example, decisions must be made on the level of detail of test scores to store for a variety of tests. Is the overall score on the test sufficient for research purposes, or should performance on individual questions or groups of questions be recorded? Is right/wrong a sufficient indicator of performance in answering a question, or is more detail needed? Similarly, in keeping records on supplementary help for students by parents or tutors, decisions must be made about the ways to indicate frequency of help, length of session, type of help (reading, drill, playing "store," or word games), and perhaps about ways to classify attitudes and their evolution. Furthermore, the relation of educational costs to effectiveness will require new breakdowns of traditional accounting information.

The way these questions are decided in turn determines the information system's cost and effectiveness. But the people making these decisions are generally quite external to the policymaking structure. If their individual efforts are to result in a consistent, efficient, and useful information system, they must be provided with well-defined guidelines concerning information privacy, the use of sociological data, the data-collection load, and other factors.

Extent to Which Information System Should be Concentrated
at a Single Center [*]

Centralized information-processing operations decrease duplication of effort and allow more economies of scale in equipment and personnel. It is also much easier to identify information-processing costs and relate them to information-processing effectiveness in a centralized operation. If the organization considering implementation is a school district, county, or state, much can be said for the benefits of centralization.

On the other hand, a more decentralized information system would generally respond more rapidly to school administrators and teaching personnel, and would also provide more local checks on the accuracy and consistency of data. A central-computer, remote-terminal system could provide some of the advantages of both centralized and decentralized operations.

INFORMATION-SYSTEM DECISIONS PRIMARILY AFFECTING INFORMATION-SYSTEM PERFORMANCE

Choices of Hardware, Information Request Monitoring, and Data Structures [***]

The information system will have a number of basic tasks to perform, including:

- o Entering standard types of data into the system.
- o Entering new types of data into the system.
- o Formatting, abbreviating, and coding of data.
- o Producing standard reports.
- o Producing new types of reports (previously unanticipated).
- o Answering specific, one-time requests.
- o Processing information (e.g., sorting and correlating).
- o Reacting to various errors.
- o Relating to other information systems.

How well the system does these jobs, and the cost it incurs, depend 1) on the choices made on the nature of the hardware and software components, and 2) on the division between manual and computer operations. For example, choosing the number of "keys" by which information is stored produces different levels of retrieval efficiency, storage costs, and flexibility. Library books are usually catalogued with three keys: author, title, and subject.

In a school system, student records maintained manually may be keyed only by name. A question such as "What students live in non-English-speaking homes?" usually requires a search through the entire file (if that information is present), even though the number of such students is probably small. A computer is able both to search the whole file quickly, and to organize it in many ways to allow direct retrieval of information without expensive searches.

The computer is particularly valuable in complex processing of filed information. "How do students' grades correlate from one year to the next?" is a reasonable question for a computer information system, an unreasonable one for a manual system. A convenient, natural way (a special "inquiry language") of expressing such requests to the computer is helpful.

Whether or Not to Build a Prototype Information System [*]

Building a prototype is a tactic available to information-system designers facing a relatively new and undefined task. A prototype allows experimentation with facets of the information-system design, without all the inertias of a full operating system. However, it requires that extra time be allocated for project definition.

* * * *

It is important to reiterate the *sequential* nature of 1) establishing information-system policies, 2) determining the accuracy, volume, and retrievability of the information to be maintained, and 3) designing and implementing the hardware-software system. As emphasized in the decision on overall time-phasing (see p. 4), an attempt to squeeze these three phases into one concurrent process usually results in a loss rather than a gain of elapsed time.

It is equally important to stress the necessity for iteration and feedback between teachers, students, administrators, research teams, and the information-system designers. Unless mechanisms exist for representatives of each of these groups to contribute to the decisionmaking processes of the other groups, the information system will lose effectiveness, and take even longer to develop.